

AD-A039 876

YALE UNIV NEW HAVEN CONN DEPT OF COMPUTER SCIENCE
SOME RESULTS ON MAXIMUM A POSTERIORI PROBABILITY PARSING ALGORI--ETC(U)
MAR 76 R J LIPTON, L SNYDER, S E LEVINSON

F/G 5/7

N00014-75-C-0752

UNCLASSIFIED

RR-98

NL

1 OF 1
ADA039 876



END

DATE
FILMED

6-77

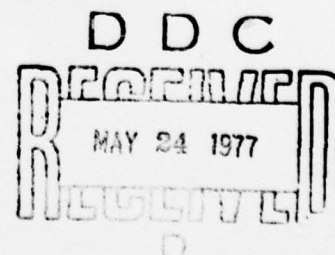
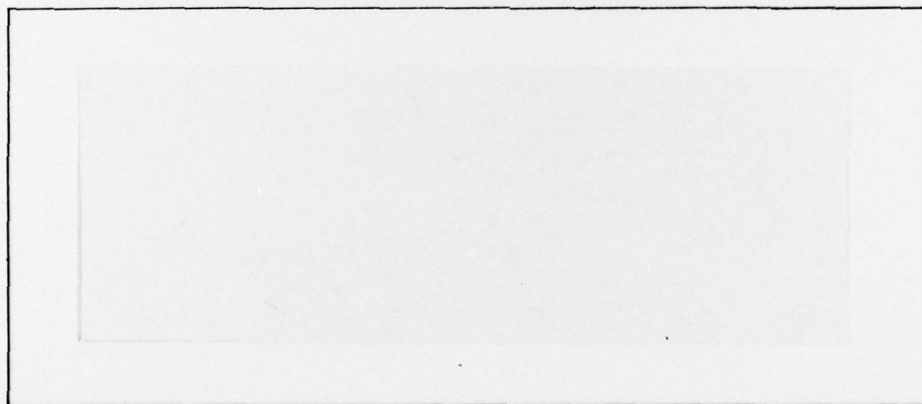
REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING PAGE
1. REPORT NUMBER RR-98 ✓	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER 12
4. TITLE (and Subtitle) 6 Some Results on Maximum a Posteriori Probability Parsing Algorithms.		5. TYPE OF REPORT & PERIOD COVERED 9 Technical Report B.S.
7. AUTHOR(s) 10 Richard J. Lipton Lawrence Snyder S. E. Levinson		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Yale University Department of Computer Science 10 Hillhouse Ave., New Haven, CT 06520		6. CONTRACT OR GRANT NUMBER(s) 15 N00014-75-C-0752 ✓
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Information Systems Program Arlington, Virginia 22217		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE 11 MAR 76 1976
		13. NUMBER OF PAGES 12 7p.
		15. SECURITY CLASS. (if this report) Unclassified
16. DISTRIBUTION STATEMENT (of this Report) Distribution of this report is unlimited		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
17. DISTRIBUTION STATEMENT (of the abstract entered in block 20, if different from Report)		DDC RECEIVED MAY 24 1977 D
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) speech parsing Viterbi Algorithm entropy		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) A fast method of parsing noisy speech is sketched. It is then used to study the entropy of languages. ↑		

AD A 039876

AD NO.

DDC FILE COPY

409 051 - B.A.M.C.



YALE UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE

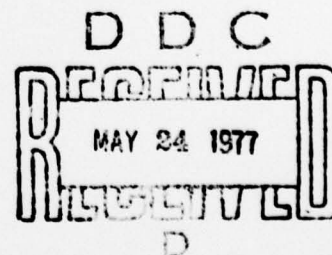
ACCESSION FOR	
NTIS	White Section <input checked="" type="checkbox"/>
DDC	Buff Section <input type="checkbox"/>
UNANNOUNCED	<input type="checkbox"/>
CLASSIFICATION	
DISTRIBUTION/AVAILABILITY CODES	
AVAIL. and/or SPECIAL	
A	

Some Results on Maximum a Posteriori
Probability Parsing Algorithms[†]

R. J. Lipton, L. Snyder
and S. E. Levinson

Research Report #98

Department of Computer Science
Yale University
New Haven, Connecticut 06520



[†] This work was supported by The Office of Naval Research
under Grant N00014-75-C-0752.

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

Some Results on Maximum a Posteriori Probability

Parsing Algorithms

R. J. Lipton, L. Snyder
and S. E. Levinson

Department of Computer Science
Yale University
New Haven, Connecticut 06511

In this paper we discuss the theory, performance and implications of a class of Syntactic Pattern Recognition algorithms which are optimal in a well defined sense.

Assume that it is desired to transmit a message consisting of a sequence of symbols chosen from a finite alphabet. Suppose further that any such message will be a well formed sentence in a language generated by a known formal grammar. The message is to be encoded and transmitted by sending a sequence of complex signals, one signal for each symbol in the message, through a noisy channel.

The corrupted message is decoded in two stages. First, the individual symbols are identified by a maximum a posteriori probability decision rule. The resulting string of symbols will not, in general, be a grammatically correct sentence. Thus, in stage two, a parser which finds that sentence in the language which maximizes the product of the individual symbol probabilities conditioned on the signals received at the output of the channel is used. The decoding thus obtained is the maximum likelihood estimator of the transmitted message.

Viterbi [1] treats the above described problem as one of estimating the state sequence of a finite Markov process. The desired estimator is obtained by a dynamic programming technique. Recently, Fung and Fu [2,3] have described an algorithm for the case of messages which are sentences in a context free language. Their procedure is based on an algorithm given by Younger [4]. We have derived efficient recursive procedures which solve the problem for regular, one-counter and context free languages. The space and time complexity of these algorithms in terms of n , the length of the input, is summarized in the table below.

$G(V_N, V_T, S, P)$	Space	Time
Regular	$ V_N n$	$ P n$
One-counter	$ V_N n^2$	$ P n^2$
Context free	$O(n^2)$	$ P n^3$

In addition to the analysis, we have tested these algorithms for several formal grammars using both real and simulated channels. Because of the efficiency of these algorithms the tests were conducted on several thousands of sentences. Some of the test grammars had over 300 production rules. The tests were accomplished without special programming considerations.

In the course of our experiments with the algorithms, we discovered an empirical measure of the information content of formal languages. By making the signal to noise ratio of the channel very low, we can reduce the performance of the single symbol decoder to the extent that it makes a random choice for each symbol independent of the input to the channel.

Although the symbol accuracy of the MAP parser also degrades with decreasing SNR, its limiting value is greater than that of the single symbol decoder alone. The difference between the two limiting values is a measure of the information encoded in the grammar.

For a given grammar $G (V_N, V_T, S, P)$ and the language, $L(G)$, it generates we define

$$L_n = \{w \in L(G) \mid |w| = n\}$$

$$V_n = \{w \in V_T^* \mid |w| = n\}$$

Then we define the entropy $H(L(G))$ of the language $L(G)$ by:

$$H(L(G)) = - \sum_n \frac{|L_n|}{|V_n|} \log_2 \left(\frac{|L_n|}{|V_n|} \right)$$

We have observed that for two grammars G_1 and G_2 , $H(L(G_1))$ and $H(L(G_2))$ are in the same order as the differences between the limiting values for their single symbol decoding and MAP parsing accuracies.

Forney [5] has listed several important problems of the type described here and has suggested that they be solved by the Viterbi Algorithm. Because this algorithm has an exponential running time it is intractible for long inputs. Forney further suggests that secondary information be used to guide a heuristic search. Such backtracking procedures have been used by Neely and White [6], Walker [7] and Levinson [8] in speech recognition algorithms.

We have observed that by applying the appropriate one of our algorithms to a variety of pattern recognition problems both the high cost of the Viterbi algorithm and the obvious disadvantages of sub-optimal backtracking procedures can be avoided and the optimal solution still obtained.

References

- [1] Viterbi, A. J., Error bounds for convolutional codes and an asymptotically optimal algorithm, IEEE Trans. Inform. Theory, IT-13, March 1967.
- [2] Fung, L. W. and Fu, K. S., Maximum likelihood syntactic decoding, IEEE Trans. Inform. Theory, IT-21, July 1975.
- [3] Fung, L. W. and Fu, K. S., Stochastic syntactic decoding for pattern classification, IEEE Trans. Comp., C-24, June 1975.
- [4] Younger, D. H., Recognition and parsing of context free languages in time n^3 , Information and Control 10(2), 1967.
- [5] Forney, G. D., The Viterbi algorithm, Proc. IEEE, 61(3), 1973.
- [6] Neely, R. B. and White, G. M., On the use of syntax in a low cost real time speech recognition system, Inf. Proc. 74, North Holland, 1974.
- [7] Walker, D. E., Speech understanding through syntactic and semantic analysis, 3rd IJCAI, SRI, 1967.
- [8] Levinson, S. E., The vocal speech understanding system, 4th IJCAI, Tbilisi, USSR, 1975.